

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2003, 19:05:45 : Search time 229 Seconds
(without alignments)
5752.514 Million cell updates/sec

Title: US-09-692-257A-1

Perfect score: 488

Sequence: 1 cctaaaaatagtttcgtata.....gaagatggcgctggcgcata 488

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

	1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:	24:	25:
	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	44.6	9.1	273	24	ABL71419
C 2	41	8.4	674	21	Corn tassal-derive 674 bp region of p
C 3	41	8.4	2796	6	AAAG61490
C 4	41	8.4	4897	11	AAAG50036
C 5	40	8.2	11396	22	AAQ03259
C 6	39.4	8.1	2742	19	ABAI7777
C 7	38	7.8	465	24	AAV49291
C 8	36.2	7.4	2064	14	ABL93562
					AAQ52638
					Corn tassal-derive 674 bp region of p
					Partial sequence o
					Pseudorabies virus
					Human nervous syst
					Aujeszky's disease
					Arabidopsis thalia
					Streptomyces fradi

C 9	36	7.4	1147	18	AAT97188
C 10	36	7.4	1906	22	AAF93877
C 11	36	7.4	25360	22	AAF88314
C 12	36	7.4	29736	22	AAF88317
C 13	35.8	7.3	3283	22	AAS63231
C 14	35.6	7.3	1473	18	AAT64548
C 15	35.6	7.3	1473	24	ABK51026
C 16	35.6	7.3	1506	24	ABK51025
C 17	35.6	7.3	1969	18	AAT64531
C 18	35.6	7.3	1969	24	ABK51009
C 19	35.6	7.3	2096	18	AAT64543
C 20	35.6	7.3	2096	24	ABK51021
C 21	35.6	7.3	2099	18	AAT64530
C 22	35.6	7.3	2099	24	ABK51008
C 23	35.6	7.3	2137	18	AAT64542
C 24	35.6	7.3	2137	24	ABK51020
C 25	35.6	7.3	2140	18	AAT64529
C 26	35.6	7.3	2140	24	ABK51007
C 27	35.6	7.3	35133	23	ABL50991
C 28	35.6	7.3	35134	23	ABL50990
C 29	35.4	7.3	417	24	ABS68109
C 30	35.4	7.3	956	24	ABL89975
C 31	35.4	7.3	80557	24	ABX09142
C 32	35.4	7.3	4403765	22	AA199683
C 33	35.4	7.3	4411529	22	AA199682
C 34	35.2	7.2	3193	20	AA191252
C 35	35	7.2	428	21	AAC56164
C 36	35	7.2	501	25	ABX64419
C 37	34.6	7.1	1662	23	AAS76807
C 38	34.6	7.1	1796	24	ABS51676
C 39	34.6	7.1	2123	21	AAC60997
C 40	34.6	7.1	2452	22	AAH17574
C 41	34.6	7.1	4799	21	AAC76258
C 42	34.4	7.0	1896	21	AAZ48297
C 43	34.4	7.0	2940	21	AAA36699
C 44	34.4	7.0	3106	21	AAA36696
C 45	34.4	7.0	3540	24	AAI70820

ALIGNMENTS

RESULT 1

ABL71419/C

ID ABL71419 standard; cDNA; 273 BP.

XX AC ABL71419;

XX AC ABL71419;

XX 14-MAY-2002 (first entry)

XX Corn tassal-derived polynucleotide (cdps) SEQ ID NO:793.

XX Corn; corn tassal-derived polynucleotide; cdps; hybrid breeding; CDPs; inheritance; characteristic; growth; development; disease resistance; environmental adaptability; quality; yield; molecular marker; multigene trait; plant breeding; corn tassal; gene; ss.

XX Zea mays.

XX Zea mays.

XX US2001051335-A1.

XX 13-DEC-2001.

XX 16-APR-1999; 99US-0294093.

XX 21-APR-1998; 98US-082567P.

XX (LALG/) LALGUDI R V.

XX (ITOL/) ITO L Y.

XX (SHER/) SHERMAN B K.

XX Laigudi RV, Ito LY, Sherman BK;

XX

PF 17-JAN-2001; 2001WO-US01334.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 23-SEP-2000; 2000US-0234997.
PR 23-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT

PA (HAMI/) HAMILTON C M.
 PA (PRIC/) PRICE J L.
 PA (RAIN/) RAINES T M.
 PA (YUYU/) YU Y.
 PA (RAME/) RAMEAKA J G.
 PA (PAGE/) PAGE A.
 PA (MATH/) MATHW A V.
 PA (LEDF/) LEDFORD B L.
 PA (WOES/) WOESSNER J P.
 PA (HAAS/) HAAS W D.
 PA (GARC/) GARCIA C A.
 PA (KRICK/) KRICKER M.
 PA (SLAT/) SLATER T.
 PA (DAVI/) DAVIS K R.
 PA (ALLE/) ALLEN K.
 PA (HOFF/) HOFFMAN N.
 PA (HURB/) HURBAN P.
 XX
 PI Goriach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
 PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
 PI Hurban P;
 XX
 DR WPI: 2002-267486/31.
 XX
 XX
 PT New Arabidopsis thaliana nucleic acid, for identifying homologous
 PT genes, producing compositions that modulate the expression or function
 PT of its encoded protein, and mapping functional regions of a protein -
 XX
 PS Claim 1: SEQ ID 327: 44pp: English.
 XX
 CC The present invention describes an Arabidopsis thaliana nucleic acid (I)
 CC comprising a sequence capable of hybridising under stringent conditions
 CC to a sequence (S1) selected from any one of the 999 sequences given in
 CC ABU93236 to ABU94234. (I) have insecticide and fungicide activities, and
 CC they can be used as protein expression modulators. (I) can be used in
 CC identifying homologous or related genes, in producing compositions that
 CC modulate the expression or function of their encoded proteins, mapping
 CC functional regions of the proteins, and in studying associated
 CC physiological pathways. (I) can also be used: (1) for the genetic
 CC manipulation of cells, particularly plant cells; (2) in screening assays
 CC of various plant strains to determine the strains that are best capable
 CC of withstanding a particular disease or environmental stress; (3) for
 CC enhancing or inhibiting production of a biosynthetic product in a plant;
 CC (4) as probes in mapping and in diagnosis, in genetic modification and
 CC for screening purposes, to generate additional copies of the nucleic
 CC acids, to generate ribozymes or antisense oligonucleotides, and as
 CC single-stranded DNA probes or as triple-stand forming oligonucleotides;
 CC and (5) for generating genetically modified transgenic organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.
 XX
 SQ Sequence 465 BP; 149 A; 101 C; 97 G; 118 T; 0 other;
 Query Match 7.8%; Score 38; DB 24; Length 465;
 Best Local Similarity 52.5%; Pred. No. 0.22; Indels 0; Gaps 0;
 Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
 QY 306 GGCAGGTGTTGGGACTGGCAGTGGCAGTGTGGCCGCCAGCGGAGAGACCT 365
 DB III IIIIII III IIIIII III IIIIII III IIIIII III IIIIII
 273 GCGCAGTGTGTATGATTTTGCAGTTTGGGCACTTCTTGAGCCACCATGTTACTGTACCT 332
 QY 366 GCGTCGTGTCCTGTCAGTGGTTCACACAGCACCATCCCTTCCCGTGTACAGTACTTG 425
 DB III IIIIII III IIIIII III IIIIII III IIIIII III IIIIII
 333 CTGCTTCTTCTCCACAGTGGTTCAGAGATCTGAACCATTTCTGTTCTGTATGTTCTG 392
 QY 426 CTTCCATCTCTGCGTGGAGGCCCTTGAAGAACTTGTCC 463
 DB III III III III III III III III III III III III III III III
 393 GCATTGGTGTAGCGGCAATCTCCATGTCAAACTTCTCC 430
 RESULT 8

AAQ52638/c
 ID AAQ52638 standard; DNA; 2064 BP.
 XX
 AC AAQ52638;
 XX
 DT 25-MAR-2003 (updated)
 DT 02-JUN-1994 (first entry)
 XX
 DE Streptomyces fradiae glutamic acid-specific protease gene.
 XX
 KW endopeptidase; glutamic acid-specific protease; Sfase;
 KW protein sequencing; ds.
 XX
 OS Streptomyces fradiae (ATCC 14544).
 XX
 FH Key Location/Qualifiers
 FT -35_signal /tag= a
 FT -10_signal 378..383
 FT CDS 435..1508
 FT /tag= c
 FT /product= glutamic_acid-specific_endopeptidase
 FT sig_peptide 435..944
 FT /tag= d
 XX
 PN W09323530-AL.
 XX
 PD 25-NOV-1993.
 XX
 PF 30-APR-1993; 93WO-JP00592.
 XX
 PR 19-MAY-1992; 92JP-0126511.
 XX
 PA (SHIO) SHIONOGI SEIYAKU KK.
 XX
 PI Kitadokoro K, Nakamura E, Shin M, Teraoka H, Tsuzuki H;
 XX WPI: 1993-386564/48.
 DR P-PSDB; AAR44216.
 XX
 PT Protease which cleaves peptide(s) at glutamic acid residue
 PT C-terminal - is isolated from Streptomyces fradiae ATCC-14544
 XX
 PS Claim 5; Page 29-32; 41pp; Japanese.
 XX
 CC This sequence encodes a protease which cleaves on the C-terminal
 CC side of glutamic acid residues and was isolated from S. fradiae ATCC
 CC 14544 genomic DNA. The protease has an optimum pH of 8.2 and a stable
 CC pH range at 37 deg C of 6-9. It can be used for the analytical
 CC investigation of proteins or to cleave peptide chains at a defined
 CC site.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 2064 BP; 270 A; 859 C; 699 G; 236 T; 0 other;
 Query Match 7.4%; Score 36.2; DB 14; Length 2064;
 Best Local Similarity 48.3%; Pred. No. 1.5;
 Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
 QY 276 GTGGCTACGAAGAGCGGGGTGCCACCGGAGGTGTGTGGGACTGGCAGTGGCAGC 335
 DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
 774 GTGCTTCAGGGGGCCATGGCGCTCGAGCTCGGGGGCGCGCTGACCCCGCGG 715
 QY 336 ACTTGTGGCCCGCCAGCGGAGAACAGCTGGTTCGTGCTGCTGCTGCTGCTGCTGCTG 395
 DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
 714 GCGTGGCGCGCGCGCGGACCTTGGCGGGCGCGGCTCGGTGTGAGGTGACGACGCA 655
 QY 396 CCCATCCCTTCCCGTGTACAGTAGCTTGCCTCCATCTCTGGTCGAGGCGCTTGAAGA 455
 DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
 654 GTTCGCGCGGTGCGCGGCTCCAGTAGTGGCGGCGCTGTCTGTCGCGGAGGTCGCTCGA 595
 QY 456 ACTTGTCCATGTGCGAAGATGGCGTGGCG 484

[illegible]

CC genome which contains the coding regions for proteins involved in
CC forosamine and trimethylrhamnose biosynthesis.
XX
SQ Sequence 29736 BP; 4401 A; 10346 C; 10080 G; 4909 T; 0 other;
Query Match 7.4%; Score 36; DB 22; Length 29736;
Best Local Similarity 48.1%; Pred. No. 5.3;
Matches 102; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 275 GGTGGTACGAGACGGGGTCCACCGGAGGTGTGTGGGACTGCGAGTGGCAG 334
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 17111 GGTGGGACATAGACGATCGGCTCCAGGTAGAGTTCCACGAGTCAGAACTGAAG 17052
QY 335 CACTGTGTGGCCCGCCACCGGAGAACCTGCGTGTGCTGCACTGCTTGACACG 394
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 17051 ATCGGACGGCCCGCCAGCGCGGCGGACAGCGCAGGCCACCCGCTAAGATGCC 16992
QY 395 ACCCATCCCTTCCCGTGTACAGTAGCTTCCCTCCATCTCTGCGTCGAGGCGCTTGAAG 454
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16991 GCTGTCGAGGCGGCTCCATCCGCGGCGCTCTCCAGCTCCGACGGCGGCGATGAAG 16932
QY 455 AACTGTGTCATGTCGAAGATGGCGTGCGCA 486
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16931 AACTGGCGCATGATGAAGATCGCCAGCACACA 16900

RESULT 13
AAS63231/c
ID AAS63231 standard; cDNA; 3283 BP.
XX
AC AAS63231;
XX
DT 29-JAN-2002 (first entry)
XX
DE DNA encoding Thermus thermophilus Muts DNA mutation binding protein.
XX
KW DNA mutation-binding protein; nuclease; DNA mismatch; cancer; PCR primer;
KW DNA damage; human xeroderma pigmentosum complementation group; XPE; XPA;
KW XPC; XPE; ERCC4; human Muts homologue 2; hMSH2; Muts; Nuc; MutY; Fpg; Ss;
KW Fapy-DNA glycosylase; uracil DNA glycosylase; ung; TDG; xthA gene; Uvr A;
KW A/G-specific adenine glycosylase; synthetic T4 endonuclease V; T4 endo V;
KW thymine DNA-glycosylase; Uvr B; Uvr C; nth gene; nfo gene; exonuclease;
endonuclease.
XX
OS Thermus thermophilus.
XX
PN WO2001173079-A2.
XX
PD 04-OCT-2001.
XX
PF 26-MAR-2001; 2001WO-US09700.
XX
PR 28-MAR-2000; 2000US-192764P.
PR 29-AUG-2000; 2000US-0650855.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Mc Cutchen-maloney SL;
XX
XX WPI; 2001-656920/75.
DR P-PSDB; AAU69744.
XX
PT Recombinant chimeric protein, useful for detecting and quantifying DNA
PT mutations, e.g. in disease diagnosis, comprises mutation-binding
PT protein and nuclease
XX
PS Claim 50; Page 73-74; 128pp; English.
XX
CC Sequences AAS63224-AAS63305 represent cDNA encoding and PCR primers used
CC to amplify cDNA encoding proteins which can be used in the synthesis of
CC chimeric proteins comprising a DNA mutation-binding protein, a linker and
CC a nuclease, by recombinant technology. The chimeric proteins are useful
CC for detection, quantification and mapping of DNA sequence variations

CC including mutations, for example, caused by damage and mismatches. The
CC proteins are able to bind to the site of the DNA mutation and cut it out
CC of the molecule. This is useful for early diagnosis of cancer and other
CC diseases. The proteins used in the invention include human XPF (or
CC ERCC4), human xeroderma pigmentosum complementation groups A, C and E
CC (XPA, XPC and XPE), human Muts homologue 2 (hMSH2), Serratia marcescens
CC nuclease (Nuc), Thermus thermophilus Muts, Escherichia coli Fapy-DNA
CC glycosylase (Fpg), uracil DNA glycosylase (ung), A/G-specific adenine
CC glycosylase (MutY), synthetic T4 endonuclease V (T4 endo V), thymine
CC DNA-glycosylase (TDG), E. coli Uvr A, B and C, and E. coli endonucleases
CC and exonucleases.
XX
SQ Sequence 3283 BP; 514 A; 1093 C; 1139 G; 537 T; 0 other;
Query Match 7.3%; Score 35.8; DB 22; Length 3283;
Best Local Similarity 52.3%; Pred. No. 2.4;
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 304 CCGGCAGGTGTGTGGGACTGGCAGTGGCAGCACTTGTGGCCCGCCAGCGGAGAAC 363
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2086 CTGCGGTCTTGTAGGCTCTGACGGGCGGTACTCTCTGGGGCACCTTCTCTAGTAGGG 2027
QY 364 CTGCGTCTGTCTCCCTGTCAGTCTGTCACACAGCCCATCCCTTCCCGTGTACAGTAGCT 423
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2026 CCGGGTCACCTCCAGGTAGTAGCCGAGACGGCGTGTACCCACCTTGGAGGTGGGAT 1967
QY 424 TGCCTCCATCTCTGCGTCGAGGCGCTTGAAG 454
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1966 GCCCGTCTCTCTCTCTCTCCCGGCGCTCGAGG 1936

RESULT 14
AAT64548/c
ID AAT64548 standard; cDNA; 1473 BP.
XX
AC AAT64548;
XX
DT 10-JUL-1997 (first entry)
XX
DE NADP-specific glutamate dehydrogenase beta subunit cDNA.
XX
KW Glutamate dehydrogenase; NADP-GDH; alga; nitrogen metabolism;
KW chloroplast; transgenic plant; ds.
XX
OS Chlorella sorokiniana strain UTEX 1230.
XX
FH Key Location/Qualifiers
FT CDS 4..1434
FT FT /*tag= a
FT mat_peptide 4..1431
FT FT /*tag= b
XX
PN WO9712983-A1.
XX
PD 10-APR-1997.
XX
PF 03-OCT-1996; 96WO-US15921.
XX
PR 06-OCT-1995; 95US-0541033.
XX
XX (UYFL) UNIV FLORIDA.
XX
PI Miller P, Schmidt RR;
XX
XX WPI; 1997-226226/20.
DR P-PSDB; AAW15412.
XX
PT DNA encoding NADP-GDH alpha and beta sub-units - isolated from
PT Chlorella sorokiniana algal protoplast(s), useful to modulate
PT nitrogen metabolism plant cells
XX
PS Claim 2; Page 48-50; 61pp; English.
XX

CC cDNA clones (AA165457 and AA165458) respectively code for the alpha
 CC (AA165411) and beta mature subunits (AA165412) of ammonium-inducible,
 CC chloroplast-localised NADP-specific glutamate dehydrogenase (NADP-
 CC GDH) hexameric isoenzymes of *Chlorella sorokiniana*. They were
 CC obtd. by removal of the chloroplast targeting signal from full-
 CC length cDNA clones (see also AA165452-43) by PCR mutagenesis. The N
 CC metabolism of plant cells can be modulated (pref. increasing the
 CC assimilation of inorganic N into organic N) by transforming them
 CC with nucleotide sequences encoding the alpha and/or beta mature
 CC subunits or subunit precursors (see also AA16529-30). Such plants
 CC show improved properties, e.g. increased crop yield and improved
 CC stress tolerance. Heterohexamers having alpha and beta subunits
 CC can be expressed that have higher aminating/deaminating activity
 CC ratios (i.e. higher capacity for glutamate synthesis) than
 CC homohexamers.

XX
 SQ Sequence 1473 BP; 297 A; 486 C; 470 G; 220 T; 0 other;
 Query Match 7.3%; Score 35.6; DB 18; Length 1473;
 Best Local Similarity 49.5%; Pred. No. 2;
 Matches 92; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 287 AGACGGGGGGTCCACCCGGCAGGTGTTGGGACTGGGAGTGGGAGCATTGTGGCCC 346
 DB 1238 AGCCCGCTGACCCGACGCGCGCGGCTTGGCCGCTTGGCGGCGAGTAGATGCCG 1179
 QY 347 GCCACGCGGAGAGACCTGCGTGTCTCCAGTCGTGCACAGCAGCCCATCCCTTT 406
 DB 1178 GCCTTGTGTACTGTGGATGGCTCGTGGTGGAGGGCAATGTGGCGCCCTCCACGACG 1119
 QY 407 CCGGTGTACAAAGTAGCTTGCCTCCATCTCTCGGTGAGGGCCCTTGAAGAACTTGTCCATG 466
 DB 1118 TACTGGCAGCGGTGCTTGATCAGCAGCTCGGCGCTGCTCATCGATCTCGTTCTGGGTG 1059
 QY 467 TCGAAG 472
 DB 1058 GCGCAG 1053

RESULT 15
 ABK51026/c
 ID ABK51026 standard; cDNA; 1473 BP.
 XX AC ABK51026;
 XX
 XX
 XX 24-SEP-2002 (first entry)
 XX cDNA encoding mature NADP-glutamate dehydrogenase beta subunit.
 DE NADP-specific glutamate dehydrogenase; NADP-GDH; beta subunit;
 KW alga; nitrogen metabolism; plant; ammonium assimilation; transgenic;
 KW ammonia toxicity tolerance; osmotic stress tolerance; gene; ss.
 XX *Chlorella sorokiniana*.
 XX
 XX Key Location/Qualifiers
 FH CDS 4..1434
 FT /*tag= a
 FT /product= "Mature NADP-specific glutamate dehydrogenase
 FT beta subunit"
 XX
 XX US2002062495-A1.
 XX
 XX 23-MAY-2002.
 XX
 XX 01-MAY-1998; 98US-0070844.
 XX
 XX 01-MAY-1998; 98US-0070844.
 XX
 XX (SCHM/) SCHMIDT R R.
 XX (MILL/) MILLER P.
 XX
 XX Schmidt RR, Miller P;

XX WPI; 2002-499691/53.
 DR P-PSDB; AAU98955.
 XX
 PT Transforming a plant with a polynucleotide encoding a polypeptide with
 PT glutamate dehydrogenase activity provides a plant with modulated
 PT nitrogen metabolism useful to increase yield and ammonium and osmotic
 PT stress tolerance
 XX
 XX Example 2; Page 31-32; 35pp; English.
 PS
 XX The invention relates to a method of modulating nitrogen metabolism in
 CC plant cells, comprising transforming a plant cell with a polynucleotide
 CC encoding a polypeptide having glutamate dehydrogenase activity, and
 CC culturing the cell to produce descendant cells which express the
 CC polypeptide. The method is used to provide plants with increased yield,
 CC improved ammonium assimilation properties, increased tolerance to
 CC ammonia toxicity, improved osmotic stress tolerance and improved
 CC composition. The present sequence represents the coding sequence of
 CC mature *Chlorella sorokiniana* NADP-glutamate dehydrogenase beta subunit.
 CC used in the method of the invention.

XX
 SQ Sequence 1473 BP; 297 A; 486 C; 470 G; 220 T; 0 other;
 Query Match 7.3%; Score 35.6; DB 24; Length 1473;
 Best Local Similarity 49.5%; Pred. No. 2;
 Matches 92; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 287 AGACGGGGGGTCCACCCGGCAGGTGTTGGGACTGGGAGTGGGAGCATTGTGGCCC 346
 DB 1238 AGCCCGCTGACCCGACGCGCGGCTTGGCCGCTTGGCGGCGAGTAGATGCCG 1179
 QY 347 GCCACGCGGAGAGACCTGCGTGTCTCCAGTCGTGCACAGCAGCCCATCCCTTT 406
 DB 1178 GCCTTGTGTACTGTGGATGGCTCGTGGTGGAGGGCAATGTGGCGCCCTCCACGACG 1119
 QY 407 CCGGTGTACAAAGTAGCTTGCCTCCATCTCTCGGTGAGGGCCCTTGAAGAACTTGTCCATG 466
 DB 1118 TACTGGCAGCGGTGCTTGATCAGCAGCTCGGCGCTGCTCATCGATCTCGTTCTGGGTG 1059
 QY 467 TCGAAG 472
 DB 1058 GCGCAG 1053

Search completed: September 12, 2003, 00:07:47
 Job time : 234 secs